Chapter 4：期望总复习

· Def： $E(X)=\sum_{j=1}^{\infty} x_j P(X=x_j)$

那么易认同：X、Y若分布相同，则 $E(X)=E(Y)$

· 关于这个定义，就可有以下 property：

① $E(X+Y)=E(X)+E(Y)$

② $E(cX)=cE(X)$

③ $P(X\geq Y)=1$，则 $E(X)\geq E(Y)$，iff $P(X=Y)=1$ 时取等

· 求期望还可用另一算法：令 $F(x)=P(X\leq x)$，令：

$$G(x)=1-F(x)=P(X>x)$$

其中，$G(x)$ 称为 survival function．则：

$$E(x)=\sum_{n=0}^{\infty} G(n).$$

Short Proof： $G(0)=P(X=1)+P(X=2)+\cdots$

$G(1)=\quad\quad P(X=2)+\cdots$

$\therefore \sum_{n=0}^{\infty} G(n)=\sum_{n=1}^{\infty} nP(X=n)$

· 求 $E(X)$ 案例：

Eg 1. $X\sim Bin(n,p)$

法1： $E(X)=\sum_{k=0}^{n} kP(X=k)=\sum_{k=0}^{n} k\binom{n}{k}p^k q^{n-k}.$

$\because k\binom{n}{k}=n\binom{n-1}{k-1}$ (story proof：n中选k-1位队员1位队长)

$\therefore \sum_{k=0}^{n} k\binom{n}{k}p^k q^{n-k}=n\sum_{k=0}^{n}\binom{n-1}{k-1}p^k q^{n-k}$

$=np\sum_{k=1}^{n}\binom{n-1}{k-1}p^{k-1}q^{n-k}=np\sum_{j=0}^{n-1}\binom{n-1}{j}q^{n-1-j}p^j$

$=np$ 明显十分复杂

法2：设X为n个相互独立且服从Bern(p)的变量之和：

$$X=I_1+\cdots+I_n.$$

$E(I_j)=p$，则 $E(X)=\sum_{i=1}^{n} E(I_j)=np$

Tip：线性性质是计算期望的一种非常有用的工具，可以避开E(X)定义

No.

Date

事实证明，这个结论十分强大！

Eg2. $X \sim HGeom(w, b, n)$ 求 $E(X)$.

解：同Eg1 法2：$X = I_1 + \cdots + I_n$ I_j代表：

取出n个中的第j个球是白球，$I_j = 1$

这里I_j之间显然不相互独立，但Linearity依然适用

那么 $P(I_j) = ?$ 呢？第j个球是白球的概率是多少？

注意不是条件概率！$P(I_j) = \frac{w}{w+b}$，i.e.:

$I_j \sim Bin(1, p)$ *

* 这不好理解，但必须意识到：在不知 $1 \sim j-1$, $j+1 \sim n$ 球是否为白的情况下，一次 I_j 就是：$w+b$ 中能否相中 w

I_j, $j \in [1, n]$ 均如此，只不过 I_j 间互有影响

$$\therefore E(X) = \frac{nw}{w+b}.$$

· LOTUS：无意识统计规律：

Def: g是 $R \to R$ 函数，则

$$E(g(X)) = \sum_x g(x) P(X=x).$$

这样，我们就不必知道 $g(X)$ 分布如何了！($g(X) = X^2$)

它能用来干什么？其中一个就是计算方差：$Var(X)$.

Def: $Var(X) = E[(X - EX)^2]$

常用：$Var(X) = E(X^2) - [E(X)]^2$

Proof: $\mu = EX$，则 $E[(X-\mu)^2] = E(X^2) - 2E(\mu X) + \mu^2$

$= E(X^2) - 2\mu EX + \mu^2 = E(X^2) - [E(X)]^2$

Var 一些 property 如下：

① $Var(X+c) = Var(X)$ ② $Var(cX) = c^2 Var(X)$

③ X, Y 独立，则 $Var(X+Y) = Var(X) + Var(Y)$

No.

Date

下面再介绍两种分布：

① 几何分布：trial 一直进行直到一次成功，令 X 表示成功前失败的次数，$X \sim Geom(p)$

PMF：$P(X=k) = q^k p$，$p+q=1$

BTW：如何判断一个PMF valid：$\sum_k P(X=k) \overset{?}{=} 1$

此处：$\sum_{k=0}^{\infty} p(q^0+q^1+\cdots) = p \cdot \frac{1}{1-q} = 1$

In other words：First Success 分布：X 代表第一次成功时的总实验次数，一定要分清 FS & Geom.

它们 Var(X) 都为 $\frac{1-p}{p^2}$

那么它的 E(X) 呢？

$E(X) = \sum_{k=0}^{\infty} kq^k p$，$q = 1-p$，看起来难以处理

观察到：$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$

则 $\left(\sum_{k=0}^{\infty} q^k\right)' = \sum_{k=0}^{\infty} kq^{k-1} = \frac{1}{(1-q)^2}$

$\therefore E(X) = pq\sum_{k=0}^{\infty} kq^{k-1} = pq\frac{1}{(1-q)^2} = \boxed{\frac{q}{p}}$.

则 FS 的 E(X) 为 $q/p + 1 = \boxed{1/p}$.

② 负二项分布：一系列 Bern(p)，X 为试验成功 r 次前失败的次数，

$X \sim NBin(r, p)$，则其 PMF 为：

$$P(X=k) = \binom{n+r-1}{r-1} p^r q^n, \quad q = 1-p$$

Theorem：$X \sim NBin(r,p)$，则 X 可表示为：

$$X = X_1 + \cdots + X_r，而 X_i \sim Geom(p)$$

Proof：共 r 次成功，每次成功前必有数次失败，故上式显然成立。

则 $E(X) = E(X_1) + \cdots + E(X_r)$，$X_i$ 独立分布.

$= \boxed{r \cdot q/p}$.

Classic Problem：Coupon Collector：

n 种 toy，欲收齐 n 种，等分布可能，求 E(N) 与 Var(N)，其中 N 代表买 toy 数量.

KOKUYO

扫描全能王 创建

No.

Date

Solution: 定义 N_i 代表买到第 i 个新种类玩具所买 toy 个数.

则 $N = N_1 + N_2 + \cdots + N_n$.

$N_1 = 1$，显然，无 toy 下，买 1 个 toy 必为新种类.

N_2 呢？成功概率为 $\frac{n-1}{n}$，故 $N_2 \sim FS(\frac{n-1}{n})$.

$\cdots\cdots$ $\therefore$ $N_i \sim FS(\frac{n-i+1}{n})$，$E(N_i) = \frac{n}{n-i+1}$

$\therefore E(N) = \sum_{i=1}^{n} E(N_i) = 1 + \frac{1}{n-1} + \frac{1}{n-2} + \cdots + n.$

$= n \cdot (1 + \frac{1}{2} + \cdots + \frac{1}{n}) \approx n(\ln n + 0.577)$. ($n$ 很大时).

- Indicator r.v. 示性随机变量

Def: $I_A = \begin{cases} 1, & \text{event } A \text{ 发生} \\ 0, & \text{event } A \text{ 未发生} \end{cases}$

则有一些基本 property:

① $(I_A)^k = I_A$ ② $I_{A^c} = 1 - I_A$ ③ $I_{A\cap B} = I_A I_B$ ④ $I_{A\cup B} = I_A + I_B - I_A I_B$

它最核心的作用，是连接 P & $E(X)$: $P(A) = E(I_A)$，且它不惧 I_j 间不独立！

这在证明一些结论有奇效：

Theorem 1: $I(A_1 \cup A_2 \cdots \cup A_n) \leq I(A_1) + \cdots + I(A_n)$

Proof: LHS = 0, ✓

LHS = 1, i.e., 至少 $\exists i$, s.t. $I_{A_i} = 1$.

则 RHS $\geq I_{A_i} = 1 =$ LHS.

Theorem 2: 容斥原理

$I(A_1 \cup \cdots \cup A_n) \overset{*}{=} 1 - I(A_1^c \cap A_2^c \cap \cdots \cap A_n^c)$.

$= 1 - (1 - I(A_1))(1 - I(A_2)) \cdots (1 - I(A_n))$

$= \sum_i I(A_i) - \sum_{i<j} I(A_i) I(A_j) + \cdots + (-1)^n I(A_1) \cdots I(A_n)$.

*: $I(A \cap B) = I(A) I(B)$，而 $I(A \cup B)$ 复杂得多.

则取补集将 $\cup \to \cap$ 是好技巧！

Campus

No.

Date

Indicator 还能用于合理化一些"直觉"

Eg: 共n人, 求生日相同的人组成的配对数的期望?

令 I_{ij}, $i<j$, $I_{ij}=\begin{cases}1, & i\text{-th}与j\text{-th}人生日相同\\0, & 不同\end{cases}$

直觉: 任调两人, 生日相同 $P=1/365$, 有 $\binom{n}{2}$ 对, 则 $E=\dfrac{\binom{n}{2}}{365}$.

现在: $E(Y)=E(I_{12}+\cdots+I_{1n}+I_{23}+\cdots)$.

$=\sum_{1\le i<j\le n} E(I_{ij}) = \binom{n}{2}\cdot P(ij) = \dfrac{\binom{n}{2}}{365}$

Indicator 还能解一些巨变态的 problem, 而它们原本变态就在于 $\boxed{I_i\text{间相互不独立}}$!

Eg. Putnam Problem: 题干略:

令 I_i 为第 i 个位置形成局部最小 $(I_i=1)$

则 $E(X)=E(I_1+\cdots+I_n)$.

而: $I_1=1/2$, 因为对于 a_1, a_2, 要么 $a_1>a_2$ 要么 $a_1<a_2$; I_n 同理.

I_i, $i\ne 1,n$ 下, $I_i=1/3$. 因为 a_{i-1}, a_i, a_{i+1}, a_i 为三数最小概率为 $1/3$.

$\therefore E(X)=2\cdot 1/2+(n-2)\cdot 1/3=\dfrac{n+1}{3}$

不敢想没 indicator, 此题能做多久时间!

Eg. 负超几何分布: X代表抽上第一个白球前, 黑球抽出数量:

注意无放回: 求 $E(X)$.

令 b个黑球标号 $1,2,\cdots,j$, I_i 代表在白球终于抽出时, i 标号黑球是否抽出.

灵性: 假设抽出白球后继续抽, 则我有许许多多的 $(w+b)$ 球抽出序列 (不标号): 再只关心 b_i 与 w 个白球, 则 b_i 在子列最前面仅一种, 共 $w+1$ 种, 则 $E(I_j)=\dfrac{1}{w+1}$.

$\therefore E(X)=\sum_{i=1}^{b} E(I_i)=\dfrac{b}{w+1}$

KOKUYO

Moment of Indicator Method:

若是$E[g(X)]$，如何用I_j理解？下面举一些例子.

① $E[\binom{X}{2}]$：可以认为是 $I_j=I_i=1$ 的pair的数量，$i<j$

i.e., $E[\binom{X}{2}] = \sum_{i<j} P(A_i\cap A_j)$，　$P(A_i)=E(I_i)$.

② $E[\binom{X}{3}]$ 呢？ $\sum_{i<j<k} P(A_i\cap A_j\cap A_k)$

③ 有了$E[\binom{X}{2}]$，则：$E(X^2) = 2\sum_{i<j}P(A_i\cap A_j)+E(X)$.

因为 $E(\frac{X^2-X}{2}) = \frac{1}{2}E(X^2) - \frac{1}{2}E(X) = \sum_{i<j}P(A_i\cap A_j)$.

接下来，再介绍一种泊淞(poisson)分布

Def：若X的PMF为：

$$P(X=k)=\frac{e^{-\lambda}\lambda^k}{k!},\quad \boxed{k=0,1,2\cdots}$$

则称X服从参数为λ的泊松分布，$X\sim Pois(\lambda)$.

Validity：$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}=e^\lambda$，故成立（$e^x$泰勒，取$x=\lambda$）.

则 $E(X)=e^{-\lambda}\sum_{k=0}^{\infty} k\frac{\lambda^k}{k!} = \lambda e^{-\lambda}\sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}$ * $=\lambda e^{-\lambda}e^{\lambda}=\lambda$.

* 无穷级数中的常见trick.　∴ $\boxed{E(X)=\lambda}$

而欲求$Var(X)$，应先知$E(X^2)$！

$E(X^2)=\sum_{k=0}^{\infty} k^2\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\lambda\boxed{\sum_{k=1}^{\infty} k\cdot\frac{\lambda^{k-1}}{(k-1)!}}$　$G(\lambda)$.

$\int G(\lambda)d\lambda = \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = \lambda e^{\lambda}$.

∴ $G(\lambda)=e^\lambda(\lambda+1)$.

∴ $E(X^2)=e^{-\lambda}\lambda e^{\lambda}(\lambda+1)=\lambda(1+\lambda)$.

$Var(X)=E(X^2)-[E(X)]^2=\lambda$　∴ $\boxed{E(X)=Var(X)=\lambda}$

No.

Date

· 泊松近似. $A_1,\cdots,A_n$, $p_j=P(A_j)$, 当 $n\to\infty$, p_j很小, A_j间独立, (n很大)

令 $X=\sum_{j=1}^n I(A_j)$. 则 X 近似于 $Pois(\lambda)$, 而 $\lambda=\sum_{j=1}^n p_j$.

用这个观点重看生日问题。生日相同pair 近似 $Pois(\lambda)$,

而 $\lambda=\binom{m}{2}\frac{1}{365}$, 则 $P(X\geq 1)=1-P(X=0)\approx 1-e^{-\lambda}$.

这个结果和之前的, 在m很大时十分近似.

Theorem: $X\sim Pois(\lambda_1)$ $Y\sim Pois(\lambda_2)$. $\boxed{X\ Y独立}$, 则 $X+Y\sim Pois(\lambda_1+\lambda_2)$

且给定 $X+Y=n$ 条件下, X条件分布: $X\sim Bin(n,\frac{\lambda_1}{\lambda_1+\lambda_2})$

Theorem: Pois 近似 Bin: (or: $\lambda\leftarrow np$)

若 $X\sim Bin(n,p)$, 令 $n\to\infty$, $p\to 0$, 且 $\lambda=np$ 固定, 则

X的 PMF 收敛于 $Pois(\lambda)$ PMF

· 如何衡量两个分布?

Def: Total Variance Distance between 分布 μ & ν (在可数集 Ω 上) is:

$$d_{TV}(\mu,\nu)=\frac{1}{2}\sum_{x\in\Omega}|\mu(x)-\nu(x)|=\|\mu-\nu\|_{TV}$$

Theorem: 小数定律:

有独立r.v. $Y_1,\cdots,Y_n$ s.t. $\forall 1\leq m\leq n$, $P(Y_m=1)=p_m$ 且

$P(Y_m=0)=1-p_m$, 令 $S_n=\sum_{i=1}^n Y_i$.

设: $\sum_{m=1}^n p_m\to\lambda\in(0,\infty)$ as $n\to\infty$.

且: $\max_{1\leq m\leq n} p_m\to 0$ as $n\to\infty$

则 $d_{TV}(S_n, Pois(\lambda))\to 0$ as $n\to\infty$.

可见: Bin 近似是上述 ~~theorem~~ 的一个例子.

大杀器: 概率生成函数 (PGF)

不关心理由, 先看定义:

PGF of 非负整数取值 r.v. X with PMF $p_k=P(X=k)$ is:

KOKUYO

No.

Date

$E(t^X)=\sum_{k=0}^{\infty} p_k t^k$ ，合理定义X，可以轻松求 p_k:

Eg: 6个 fair dice, $X_1,\cdots,X_6$ 是独立投掷. 则 $P(X=18)$?

Solution: $X=X_1+\cdots+X_6$

则 $E(t^X)=E(t^{X_1+\cdots+X_6})=E(t^{X_1})\cdots E(t^{X_6})=[E(t^{X_1})]^6$.

而 $E(t^{X_1})$ 　　$E(g(X))=\sum_x g(x)P(X=x)$

$=\sum_{x=1}^{6} t^x P(X=x)=\frac{1}{6}(t+t^2+\cdots+t^6)$

$\therefore E(t^X)=\frac{1}{6^6}t^6\cdot(1+\cdots+t^5)^6$，展开求 t^{18} 前系数即可.

而更6的结论在于: $g(t)=E(t^X)=\sum_{k=0}^{\infty}p_k t^k$，则有:

$$\begin{cases} E(X)=g'(t)\big|_{t=1} \\ E(X(X-1))=g''(t)\big|_{t=1} \end{cases}$$

Proof: $g'(t)=\sum_{k=0}^{\infty}k t^{k-1}p_k$. $g'(t)\big|_{t=1}=\sum_{k=0}^{\infty}kp_k=E(X)$.

$g''(t)=\sum_{k=0}^{\infty}k(k-1)t^{k-2}p_k$. $g''(t)\big|_{t=1}=\sum_{k=1}^{\infty}k(k-1)p_k$

$E(X-1)=\sum_{k=1}^{\infty}(k-1)p_k$ $\therefore E(X)E(X-1)=E(X(X-1))=g''(t)\big|_{t=1}$.

Eg: Pattern Matching Problem:

一个 coin, p 机率 head，不断 toss 直至有"HH"；N 为观察到"HH"时 toss 的次数，求 E(N)　　($q=1-p$).

Solution: $p_k=P(X=k)$. $p_0=0$, $p_1=0$, $p_2=p^2$, $p_3=p^2q$.

对于 p_k，若1,2次为HT，则 (k-2)次最后见HH　　$\therefore p_k=pqp_{k-2}+qp_{k-1}$

若1次为T，则 (k-1)次最后见HH　　特征根会算爆!

$E(t^N)=\sum_{k=0}^{\infty}p_k t^k=\sum_{k=1}^{\infty}p_k t^k=p^2t^2+qt\sum_{k=3}^{\infty}p_{k-1}t^{k-1}+pqt^2\sum_{k=3}^{\infty}p_{k-2}t^{k-2}$

$\therefore g(t)-p^2t^2=qt\cdot g(t)+pqt^2\cdot g(t)$.

$\therefore g(t)=\frac{p^2t^2}{1-qt-pqt^2}$. 则 $E(N)=g'(t)\big|_{t=1}$, $E(N^2-N)=g''(t)\big|_{t=1}$

$Var(N)=E(N^2)-E(N)^2$ 代入即可

核心: 递推式亦可用于求 $g(t)=E(t^X)$. 而 $E(X)$. $Var(X)$. $E(X^2-X)$ 与 $g(t)$ 挂钩!

注意递归式是如何得到的！实际规则可能复杂的多！

我总结了"鱼骨图"以精确得到递归式：($H: p$. $T: q$, $p+q=1$)

Eg："HTHT" step 1：作出：— H — T — H — T

step 2：交错补图：

H T H T
T H T H

step3：画出 focus：（所有的补充 H/T）

H T H T
T 1 H 2 T 3 H 4.

step 4：对于每一个 focus，如果它不是"HTHT"开头，i.e. H

则为"P_{k-i}"型；否则为"P_{k-i}｜第一个为H"型：（条件概率）

*：它的计算可用金-补进行

① T，则为 $q P_{k-1}$

② H，则为：$p(P_{k-1}-qP_{k-2})$*

③ T，则为：$pq^2 P_{k-3}$ $\therefore P_k=①+②+③+④$

④. H，则为：$p^2q(P_{k-3}-qP_{k-4})$.

P_{k-i} 型：$\prod\limits_{j=1}^{i} P_j \cdot P_{k-i}$. $P_j=p$ or q

P_{k-i}｜第一个为XX型：$\prod\limits_{j=1}^{i-1} P_j \cdot (P_{k-i+1}-p \text{ or } q\, P_{k-i})$.

如 HH： — H — H

T T

$pq\cdot P_{k-2}$

$q\cdot P_{k-1}$ $\therefore P_k=pqP_{k-2}+q\cdot P_{k-1}$.

No.

Date

重要分布的期望 方差一览

1. $X \sim Bern(p)$. $\Rightarrow E(X) = 1\cdot p + 0\cdot q = p$

$E(X^2) = 1^2 \cdot p + 0 \cdot q = p$ $\Rightarrow Var(X) = E(X^2) - [E(X)]^2 = p - p^2$

2. $X \sim Bin(n, p)$. 因为每一次Bern试验之间独立，记为X_i, $i \in [1, n]$

故因Linearity: $\Rightarrow E(X) = \sum_{i=1}^{n} E(X_i) = np$

二项分布 $\Rightarrow Var(X) = \sum_{i=1}^{n} Var(X_i) = np(1-p)$.

3. 超几何分布: $X \sim HGeom(n, w, b)$，令 $X = I_1 + \cdots + I_n$

I_j为1 if 第j个球为白，否则为0。则对于I_j们来说：

$p = \frac{w}{w+b}$，且相互独立

这是因为此处为I_j分好球后一齐揭开答案，而不是分一个揭一个！"

可见：关于它 $E(X)$ $Var(X)$的分析，如同二项分布：

$\Rightarrow E(X) = n \cdot \frac{w}{w+b}$

$\Rightarrow Var(X) = n \cdot \frac{w \cdot b(w+b-n)}{(w+b)^2(w+b-1)}$

$\left(E(X(X-1)) = \sum_{x=2}^{n} x(x-1) \frac{\binom{w}{x}\binom{b}{n-x}}{\binom{w+b}{n}} \right)$

4. 几何分布 & First Success: $\Rightarrow \begin{cases} E(X) = q/p & (几何) \\ E(X) = 1/p & (FS) \end{cases}$

而: $Var(X)_{几何} = Var(X)_{FS}$,

用 $E(X^2)$ 计算 即可(较复杂，此处略) $\Rightarrow Var(X) = \frac{q}{p^2}$ (几何&FS)

5. 负二项分布: 从一次成功到下一次成功间失败次数可视为$X_i \sim Geom(p)$

$\Rightarrow E(X) = r \cdot q/p$ $\Rightarrow Var(X) = r \cdot q/p^2$

6. 泊松分布: $\Rightarrow E(X) = \lambda$

$\Rightarrow E(X^2) = \lambda$